

Cornell ISS Networks Workshop:

Search and Diffusion on Social Networks  
November 8-11, 2006  
423 ILR Conference Center

**Abstracts**

*Wednesday, Nov. 8*

*Behavioral Graph Coloring*

[Michael Kearns](#), Computer and Information Science, University of Pennsylvania

The pioneering work of Travers and Milgram in 1969 established the now-familiar folklore of "six degrees" of separation in natural social networks. More recently, researchers including Jon Kleinberg and Duncan Watts have explored the algorithmic aspects of how messages are forwarded in such networks. Perhaps the computer science view of this fascinating line of thought can be best summarized as follows: Using relatively local information, distributed human organizations can compute good approximations to the all-pairs shortest paths problem. What other sorts of distributed optimization problems can humans networks solve?

In this talk, I will describe the preliminary but thought-provoking findings of a series of behavioral experiments we have been performing at Penn. Human subjects attempt to perform distributed graph coloring using a system that controls network structure, information conditions, incentives, and a variety of other variables of interest. The experiments shed early light on whether such problems can be solved by human networks, under what conditions, and on what algorithms they seem to adopt.

*Ayatollahs on the Pareto Frontier*

David Patel, Government, Cornell University

Why did a cohesive national Shiite political identity emerge in post-Baathist Iraq, but Sunni Arabs remain fractured? Why did Grand Ayatollah Ali al-Sistani and Moqtada al-Sadr surprisingly become Iraq's most important political entrepreneurs, instead of party leaders or other elites? What explains variation in parties' electoral success across and within Iraqi governates? Iraqi Arabs have needed to coordinate with each other to collectively act post-invasion, both locally and nationally, which requires a particular kind of shared information, common knowledge. Iraqis need to form shared expectations about how other Iraqis will behave in a wide-range of previously unforeseen circumstances. I argue Muslim religious leaders control the most natural way for Iraqi Arabs to get this sort of information; Friday mosque sermons create common knowledge, knowledge of others' knowledge, within geographically-defined congregations. Shiite Ayatollahs, controlling hierarchical networks of clerical deputies, can reliably and consistently disseminate similar messages in different mosques, generating common knowledge and coordination across dispersed Shiite congregations on national-level issues like federalism and voting strategies. Through mosque networks, Shiites reliably know what Shiites in far distant areas know. Sistani and Sadr control the two largest such networks. Sunni clerics lack the ability to consistently coordinate sermons across mosques, hindering the emergence of a cohesive national Sunni Arab identity and

leadership. Other post-Baathist political organizations, such as tribes, parties, the coalition authority, and media outlets, lacked sufficient ability to generate the wide common knowledge necessary to coordinate Iraqi Arabs across localities. The influence of religious authorities is not just due to the message of Islam; it is due to the ability of Islam to deliver messages. Observable implications of the theory and alternative explanations are tested on electoral data from national elections in January and December 2005.

*Cyberinfrastructure and Social Network Research*

[Noshir Contractor](#), Speech Communication, Psychology, and Coordinated Science Laboratory, University of Illinois

***Thursday, Nov. 9***

*Learning and Exploiting Statistical Dependencies in Networks*

[David Jensen](#), Computer Science, University of Massachusetts

We propose a new algorithm for finding a target node in a network whose topology is known only locally. We formulate this task as a problem of decision making under uncertainty and use the statistical properties of the graph to guide this decision. This formulation uses the homophily and degree structure of the network simultaneously, differentiating our algorithm from those previously proposed in the literature. Because homophily and degree disparity are characteristics frequently observed in real-world networks, the algorithm we propose is applicable to a wide variety of networks, including two families that have received much recent attention: small-world and scale-free networks.

*Perspectives on Node Centrality*

[Steve Borgatti](#), Organization Studies, Boston College

Node centrality is often viewed as "just a measure" or, at the other extreme, rather fancifully in such terms as prestige, importance, status, power, independence, brokerage, and so on. This paper treads a middle course and presents both a graph-theoretic view and a dynamic process view of centrality.

*Public Goods in Networks*

[Rachel Kranton](#), Economics, University of Maryland

This paper examines experimentation and innovation when experimental results diffuse along social links. This public-goods nature of experimentation is a feature of many areas of economics, such as consumer choice, research and development, and agriculture. The paper asks: How do different patterns of social and geographic links affect experimentation? Who experiments and who free rides? Do more links enhance or diminish social welfare? The analysis finds, first, that social networks can foster specialization. In every social network there is an equilibrium where some individuals experiment and others completely free-ride. In many networks this extreme is the only equilibrium outcome. Second, specialization can benefit society as a whole. This outcome arises when specialists are linked, collectively, to many agents. Finally, new links can

reduce overall welfare. A new link increases access to new results, but also reduces an individual's incentives to conduct his own experiments. Hence, overall welfare can be higher when there are holes in a network.

*Algorithmic Models for Search and Contagion*

[Jon Kleinberg](#), Computer Science, Cornell University

The processes by which communities come together, attract new members, and develop over time is a central research issue in the social sciences—political movements, professional organizations, and religious denominations all provide fundamental examples of such communities. In the digital domain, on-line groups are becoming increasingly prominent due to the growth of community and social networking sites such as MySpace and LiveJournal. However, the challenge of collecting and analyzing large-scale timeresolved data on social groups and communities has left most basic questions about the evolution of such groups largely unresolved: what are the structural features that influence whether individuals will join communities, which communities will grow rapidly, and how do the overlaps among pairs of communities change over time? Here we address these questions using two large sources of data: friendship links and community membership on LiveJournal, and co-authorship and conference publications in DBLP. Both of these datasets provide explicit user-defined communities, where conferences serve as proxies for communities in DBLP. We study how the evolution of these communities relates to properties such as the structure of the underlying social networks. We find that the propensity of individuals to join communities, and of communities to grow rapidly, depends in subtle ways on the underlying network structure. For example, the tendency of an individual to join a community is influenced not just by the number of friends he or she has within the community, but also crucially by how those friends are connected to one another. We use decision-tree techniques to identify the most significant structural determinants of these properties. We also develop a novel methodology for measuring movement of individuals between communities, and show how such movements are closely aligned with changes in the topics of interest within the communities.

*Diffusion of Behavior and Equilibrium Properties in Network Games*

[Matthew Jackson](#), Political Science, Stanford University

We analyze games on social networks where agents select one of two actions (whether or not to adopt a new technology, withdraw money from the bank, become politically active, etc.). Agents' payoffs from each of the two actions depend on how many neighbors she has, the distribution of actions among her neighbors, and a possibly idiosyncratic cost for each of the actions. We analyze the diffusion of behavior when in each period agents choose a best response to last period's behavior. We characterize how the equilibrium points of such a process and their stability depend on the network architecture, the distribution of costs, and the payoff structure. We also illustrate how the dynamics of behavior depends on the number of neighbors that agents have. Our results have implications and applications to marketing, epidemiology, financial contagions, and technology adoption.

***Friday, Nov. 10***

*Identification Problems in Social Learning*

[Charles Manski](#), Economics, Northwestern University

I analyze social interactions that stem from the successive endeavors of new cohorts of heterogeneous decision makers to learn from the experiences of past cohorts. A dynamic process of information accumulation and decision making occurs as the members of each cohort observe the experiences of earlier ones, and then make choices that yield experiences observable by future cohorts. Decision makers face the selection problem as they seek to learn from observation of past actions and outcomes, while not observing the counterfactual outcomes that would have occurred had other actions been chosen. Assuming that all cohorts face the same outcome distributions, I show that social learning is a process of sequential reduction in ambiguity. The specific nature of this process, and its terminal state, depend critically on how decision makers make choices under ambiguity. I use the problem of learning about innovations to illustrate.

*Dynamic Networks, Statistics, and Simulations*

[Bruce Kogut](#), Strategy, INSEAD

The recorded transactions of venture capital investments permit a direct examination of the Braudel hypothesis that regional markets evolve dynamically and *interdependently* in reference to a global system. This hypothesis contradicts the popular belief that regional financial development is anchored in dense clusters. Using methods of complex graphs, we analyze 159,561 transactions over nearly 45 years to demonstrate the rapid emergence of a national network of syndications. A giant component emerges early in the history of the industry which subsumes the regional and sectoral sub-graphs. The results confirm the Braudel hypothesis over the role of regional clusters, rejects preferential attachment in favor of repeated ties among trusted partners, and emphasizes the importance of dynamics and complex weighted graphs for the analysis of social and economic behavior.

*Understanding the Inputs into Innovation*

[Shane Greenstein](#), Management and Strategy, Kellogg School of Management

We examine whether there is a tradeoff between employing internal (firm) resources and purchased external (local) resources in process innovation. We draw on a rich data set of Internet investments by 86,879 U.S. establishments to examine decisions to invest in advanced Internet technology. We find evidence of localization of substitution. In particular, we show that the marginal contribution of internal resources is greater outside of a major urban area than inside one. Agglomeration is therefore less important for highly capable firms. When firms invest in innovative processes they act as if resources available in cities are partial substitutes for both establishment-level and firm-level internal resources.

*When Does Networking Work?*

[Duncan Watts](#), Sociology, Columbia University

Individuals often engage in social networking in order to gather or mobilize resources distributed among other individuals. The term “networking,” however, can refer to at least two quite distinct strategies: “interpersonal networking,” in which individuals query existing acquaintances for referrals to others; or “targeted” networking, in which individuals attend events or join new groups in order to develop new ties without the aid of intermediaries. In this paper, we compare the performance of interpersonal and targeted networking strategies across a space of artificial social networks whose properties are consistent with existing sociological theory, as well as known empirical examples. We explore the conditions under which these networking strategies are effective and determine when each strategy is preferred over the other. Our results suggest that homophily, the tendency for people to associate with others like themselves is an essential property of social networks for either networking strategy to be effective, but that each strategy is preferred under a different portion of this domain.

*Diffusion of Recommendations in Online Social Networks*

[Lada Adamic](#), School of Information, University of Michigan

We present an analysis of a person-to-person recommendation network, consisting of 4 million people who made 16 million recommendations on half a million products. We observe the propagation of recommendations and the cascade sizes, which we explain by a simple stochastic model. We analyze how user behavior varies within user communities defined by a recommendation network. Product recommendations follow a long tail where a significant share of recommendations correspond to rarely sold items. We establish how the recommendation network grows over time and how effective it is from the viewpoint of the sender and receiver of the recommendations. While on average recommendations are not very effective at inducing purchases and do not spread very far, we present a model that successfully identifies communities, product and pricing categories for which viral marketing seems to be very effective.

*Second-Hand Brokerage*

[Ronald Burt](#), Sociology and Strategy, University of Chicago Graduate School of Business

The social capital of brokerage is evident from the higher compensation, more positive recognition, and broader responsibility given to people who coordinate across the structural holes between groups. This paper is about brokerage among direct versus indirect contacts. Information moved between direct contacts I discuss as direct brokerage, to distinguish it from information moved between friends of friends - people to whom one is only connected indirectly - which I discuss as second-hand brokerage. Analyzing network associations with performance in three study populations, I find that second-hand brokerage has little or no value in a wide variety of circumstances. Brokerage benefits are dramatically concentrated in the immediate network around a person. Why that is so, and conditions under which it is more or less so, are the subjects of this paper. The implication for research design is that brokerage can be measured using designs in which data are limited to the immediate network around an individual. The theory implication is that the social capital of brokerage is a local phenomenon as in the Austrian market metaphor with its emphasis on tacit knowledge about local norms

and practice. I close with next steps for future research on network metrics for organizations and micro mechanisms given new significance as success factors in the competitive advantage that research has traced to brokerage.

*Saturday, Nov. 11*

*Markov Logic: A Rich Modeling Language for Social Networks*

[Pedro Domingos](#), Computer Science and Engineering, University of Washington

Real social networks contain multiple types of actors related in multiple ways, and complex dependencies among these relations. We need a modeling language that makes it easy to express these, learn them from data, perform inference with them, and refine them. We propose Markov logic as such a language. A model in Markov logic is a set of weighted formulas in first-order logic, where each formula is interpreted as a template for features of a Markov random field, with the corresponding weight. Standard models used in social network analysis are easily expressed in Markov logic, as are much more complex ones. We have developed highly efficient algorithms for Markov logic drawing on ideas from satisfiability testing and Markov chain Monte Carlo, and learning algorithms based on the voted perceptron, pseudo-likelihood and inductive logic programming. Markov logic has been successfully applied to a variety of problems in social network modeling, and is the basis of the open-source Alchemy system.

*Limits of Diffusion Potential on Dynamic Networks*

[James Moody](#), Sociology, Duke University

This presentation explores the fundamental variability in diffusion over networks when edges are dynamic. The paper uses simulation methods to explore how edge duration and local sequences interact with network structure to propel diffusion across seemingly connected networks.